Proyecciones Journal of Mathematics
Vol. 43, N^o 1, pp. 189-224, February 2024.
Universidad Católica del Norte
Antofagasta - Chile

10.22199/issn.0717-6279-5984

Optimal modeling of nonlinear systems: Method of variable injections

Anatoli Torokhti
University of South Australia, Australia
and
Pablo Soto-Quiros
Instituto Tecnológico de Costa Rica, Costa Rica
Received : April 2023. Accepted : October 2023

Abstract

Our work addresses a development and justification of the new approach to the modeling of nonlinear systems. Let $\mathcal{F}$ be an unknown input-output map of the system with a random input and output $\mathbf{Y}$ and $\mathbf{X}$, respectively. It is assumed that $\mathbf{Y}$ and $\mathbf{X}$ are available and covariance matrices formed from $\mathbf{Y}$ and $\mathbf{X}$ are known. We determine a model of $\mathcal{F}$ so that an associated error is minimized. To this end, the model $\mathcal{T}_p$ is constructed as a sum of $p+1$ particular parts, in the form $\mathcal{T}_p(\mathbf{Y}) = \sum_{j=0}^{p} G_j H_j Q_j(\mathbf{V}_j)$ where G_j and H_j, for $j = 0, \ldots, p$, are matrices to be determined, and $\mathbf{V}_j$, for $j = 1, \ldots, p$, is a special random vector called the injection. We denote $\mathbf{V}_0 = \mathbf{Y}$. Further, Q_j is a special transform aimed to facilitate the numerical realization of model $\mathcal{T}_p$. It is determined in the way allowing us to optimally determine G_j and H_j as a solution of $p+1$ separate error minimization problems which are simpler than the original minimization problem. The empirical determination of injections $\mathbf{V}_1, \ldots, \mathbf{V}_p$ is considered. The proposed method has several degrees of freedom to diminish the associated error. They are 'degree' p of $\mathcal{T}_p$, choice of matrices $G_0, H_0, \ldots, G_p, H_p$, dimensions of matrices $G_0, H_0, \ldots, G_p, H_p$ and injections $\mathbf{V}_1, \ldots, \mathbf{V}_p$, respectively. In particular, it is shown that a variation of the injections in their dimensionality and special forms allow us to increase accuracy of the proposed model $\mathcal{T}_p$. The proposed approach differs from known techniques by its ingredients mentioned above. Four numerical examples are provided. At the end, the open problem is formulated.

Subjclass [2020]: *15A29, 65F55, 60E05, 65K10.*

Keywords: *Nonlinear systems; Error minimization; Optimization.*

1. Introduction

We propose and justify a new method for the optimal modeling of systems transforming random signals.

1.1. Motivation

The problem of a constructive representation of nonlinear systems has been a topic of profound research for a long time. A number of fundamental papers have appeared which established significant advances in this research area. Some relevant references can be found, in particular, in [31, 19, 67, 69, 68, 34, 24, 15, 22, 21, 72, 6, 20, 63, 50, 42].

The known related results mainly concern proving the existence and uniqueness of mathematical models used for an approximate representation of a system under consideration, and for justifying the bounds of errors arising from the approximation methods. The assumptions are that inputs and outputs are deterministic and can be represented in an analytical form, that is, by equations. At the same time, in many applications, the inputs and outputs are stochastic and cannot be described by equations. Nevertheless, it is possible to represent these sets in terms of their numerical characteristics, such as the expectation and covariance matrices. Typical examples are engineering [66, 44, 23, 8], statistics [15, 38, 16, 74], stochastic signal processing [61, 36, 92, 88, 9], and image processing [17, 60]; in the latter case, a digitized image, presented by a matrix, is often interpreted as the sample of a stochastic signal.

While the theory of a system representation with any given accuracy is well elaborated (see, e.g., [31, 19, 67, 69, 68, 34, 24]), the theory of *optimal constrained* and *constructive* system representation is still not so well developed, although this is an area of intensive research (see, e.g., [22, 21]). Despite increasing demands from applications [72, 6, 20, 63, 92, 88, 9, 38, 16, 74, 66, 44, 23, 8, 60, 22, 21, 28, 79, 51, 3, 32, 29, 5, 62, 65] this subject is hardly tractable because of intrinsic difficulties in optimal approximation techniques, especially when the approximating model should have a specific structure implied by the underlying problem.

We wish to extend the known results in this area to the case when the inputs and outputs of the system are stochastic, and the approximating model we search is constructive. The latter means the model can numerically be realized and, therefore, is applicable to problems in applications.

1.2. Differences from known techniques

For the case when the dimensionality reductions of signals are not required, the following results related to the problem under consideration are as follows.

The well-known Levinson-Schur algorithms (see, e.g. [78, 77]) are fast but they are only applicable to Toeplitz matrices. The covariance matrices we consider are not Toeplitz. It is possible to replace the covariance matrix with a Toeplitz matrix that is closest to a given covariance matrix [26], and then apply the Levinson-Schur algorithms to that. But this replacement will, of course, increase the associated error.

The alternating direction method of multipliers (ADMM) is well suited to the distributed convex optimization [13] (see pp. 2, 33-37) but the problem we consider is not convex. For non-convex problems, ADMM can converge to non-optimal points—see pp. 73-77 in [13]. In [57], the proposed scenario requires an use of block diagonal matrices and the existence of non-singular matrices. In our solution, we do not limit ourselves with those restrictions. In [14], a fast low-rank modifications of the thin SVD is studied, but the SVD of the covariance matrices we consider is not thin. Methods of a fast computation of the pseudo-inverse matrix in [18, 7, 55] are approximate and iterative and, therefore, the associated computational load increases with the increase in the number of iterations. The proposed method avoids iterations for the determination of the optimal system model.

We also note that the problems we consider are different from those studied in [86, 54, 41].

The optimal dimensionality reduction of a random signal is achieved by the Karhunen-Loève transform (KLT) which is studied in a number of works. We cite just a few of them in [70, 35, 36]. The KLT is also known as Principal Component Analysis [37]. Recall, the KLT is represented by a matrix of fixed rank that minimizes the error associated with the optimal compression and estimation of the source random vector s. We consider a case where the system model consists of $p + 1$ blocks transforming input signal $\mathbf{y}$. If the number of blocks $p+1$ is greater than one then the compression of data arrays cannot be done by the KLT. Therefore, to find a model of the system, for $p > 1$, a new approach should be used. In particular, for the case of two system blocks, this problem has been solved in [83].

For an arbitrary number of the system blocks, in [27, 25, 75, 93, 71, 4, 48, 45, 56, 89, 91], approximate iterative solutions of this problem are considered. They are mainly based on implementations of the block coordinate descent method (BCDM) and its modifications [84, 10]. Unfortunately, con-

vergence of the BCDM to a global minimum is derived under some heavy restrictions [84, 10] which are difficult to implement and often cannot be satisfied in practice.

The proposed method is based on an approach which is different from those mentioned above. This allows us avoid associated difficulties.

1.3. Short description of the method

Let Ω be a set of outcomes in probability space (Ω, Σ, μ) for which Σ is a σ–field of measurable subsets of Ω and $\mu : \Sigma \to [0, 1]$ is an associated probability measure. Let $\mathbf{x} \in L^2(\Omega, \mathbf{R}^m)$ and $\mathbf{y} \in L^2(\Omega, \mathbf{R}^n)$ be random input and output, respectively, such that $\mathbf{x} = \mathcal{F}(\mathbf{y})$ where $\mathcal{F} : L^2(\Omega, \mathbf{R}^n) \to L^2(\Omega, \mathbf{R}^m)$ is an unknown non-linear input-output map. It is assumed that the only available information on $\mathcal{F}$ is given by certain covariance matrices formed from $\mathbf{x}$ and $\mathbf{y}$. This is a typical assumption used in the applications such as those considered, e.g., in [61, 36, 92, 88, 9, 15, 38, 16, 74, 72, 6, 5, 62, 65]. Here, we adopt that assumption. It is also assumed that $\mathbf{x}$ and $\mathbf{y}$ are available.

We develop a new approach to the optimal constructive representation of the nonlinear system $\mathcal{F}$ subject to a specialized criterion associated with the dimensionality reduction of the random input. The latter constraint follows from the requirements in applications such as those considered in [15, 38, 16, 74, 44, 3, 32]. In particular, a dimensionality reduction of random signals is used to optimize the cost of signal transmission.

In Section 2.3 that follows, we consider a problem that concerns finding the optimal system model that depends on $2p + 2$ unknown matrices G_j and H_j, for $j = 0, \ldots, p$ and a given non-negative integer p. The difficulty is that $2p+2$ unknown matrices should be determined from a minimization of the single cost function represented in (2.6). The solution is provided in Section 3 and is given by the optimal approximating operator minimizing the associated error.

Numerical examples that illustrate the advantages of the proposed method are provided. Computations were performed on a desktop computer with an Intel(R) Core(TM) i9-10900F CPU (2.80 GHz) and 32.00 RAM using Matlab R2021a. The related code is available in GitHub. The link is https://github.com/jusotoTEC/variableInjections.

1.4. Novelty

The proposed system model $\mathcal{T}_p$ is represented by a sum of $p+1$ terms where jth term, for $j = 0, \ldots, p$, is a model of jth block of the system (see (2.5) in Section 2.2). Each term of model $\mathcal{T}_p$ contains a special transformation Q_j, for $j = 0, \ldots, p$, which leads to a faster numerical realization of the proposed method. This is because the transformations provide a representation of the problem under consideration in (2.6) by a set of simpler problem each of which depends on the single pair G_j, H_j, for $j = 0, \ldots, p$. It allows us to avoid numerical difficulties associated with computation of large matrices and determine G_j, H_j, for $j = 0, \ldots, p$ from the problem which is simpler than that in (2.6). Details are given in Section 3.2 that follows.

The model $\mathcal{T}_p$ has several degrees of freedom to minimize the associated error. They are: 'degree' p of $\mathcal{T}_p$, choices of matrices $G_0, H_0, \ldots, G_p, H_p$ (Theorems 3, 4, 5) and injections $\mathbf{v}_1, \ldots, \mathbf{v}_p$ (Section 3.3, Theorem 5, Remarks 2, 3 and 6), and dimensions $r_0, \ldots, r_p$ (Theorems 3). In particular, an influence of variations of the injections $\mathbf{v}_1, \ldots, \mathbf{v}_p$ on the decrease of the associated error is shown in Theorem 4, and Remarks 2, 3 and 6: an increase in the dimensions and of the injections and their choice improve the accuracy of the system model.

2. The proposed approach

2.1. Some special notation

Let us write $\mathbf{x} = [\mathbf{x}_{(1)}, \ldots, \mathbf{x}_{(m)}]^T$ and $\mathbf{y} = [\mathbf{y}_{(1)}, \ldots, \mathbf{y}_{(n)}]^T$ where $\mathbf{x}_{(i)}, \mathbf{y}_{(j)} \in L^2(\Omega, \mathbf{R})$, for $i = 1, \ldots, m$ and $j = 1, \ldots, n$, and $\mathbf{x}(\omega) \in R^m$ and $\mathbf{y}(\omega) \in R^n$ for all $\omega \in \Omega$.

Each matrix $A \in R^{m \times n}$ defines a bounded linear transformation $\mathcal{A} : L^2(\Omega, \mathbf{R}^n) \to L^2(\Omega, \mathbf{R}^m)$. It is customary to write A rather then $\mathcal{A}$ since $[\mathcal{A}(\mathbf{x})](\omega) = A[\mathbf{x}(\omega)]$, for each $\omega \in \Omega$.

Let us also denote

$$\|\mathbf{x}\|_\Omega^2 = \int_\Omega \sum_{j=1}^{m} [\mathbf{x}_j(\omega)]^2 d\mu(\omega) < \infty. \tag{2.1}$$

The covariance matrix formed from $\mathbf{x}$ and $\mathbf{y}$ is denoted by

$$E_{xy} = \left\{ \int_\Omega \mathbf{x}_{(i)}(\omega) \mathbf{y}_{(j)}(\omega) d\mu(\omega) \right\}_{i,j=1}^{m,n}. \tag{2.2}$$

The Moore-Penrose pseudo-inverse [11] of matrix M is denoted by $M^\dagger$.

2.2. Generic structure of the system model

Let $\mathbf{y}$ be an input signal and $\mathbf{v}_1, \ldots, \mathbf{v}_p$ be random vectors such that $\mathbf{v}_j \in L^2(\Omega, \mathbf{R}^{q_j})$, for $j = 1, \ldots, p$. We write $\mathbf{y} = \mathbf{v}_0$ and $q_0 = n$. Each $\mathbf{v}_j$, for $j = 1, \ldots, p$, is defined by a nonlinear transformation φ_j of $\mathbf{y}$, i.e., $\mathbf{v}_j = \varphi_j(\mathbf{y})$. The choice of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ is considered in Section 3.3. To facilitate the numerical implementation of the system model introduced below, each vector $\mathbf{v}_j$, for $j = 1, \ldots, p$, is transformed to vector $\mathbf{z}_j \in L^2(\Omega, \mathbf{R}^{q_j})$ by transformation Q_j so that

$$\mathbf{z}_j = Q_j(\mathbf{v}_j, Z_{j-1}), \tag{2.3}$$

where $Z_{j-1} = \{\mathbf{z}_0, \ldots, \mathbf{z}_{j-1}\}$. The choice of Q_j is considered in Section 3.1. We call $\mathbf{v}_1, \ldots, \mathbf{v}_p$ the injections. This is because $\mathbf{v}_1, \ldots, \mathbf{v}_p$ contribute to the decrease of the associated error as shown in Section 3.4 below. At the same time, in particular, $\mathbf{v}_j$ can be a noisy version of $\mathbf{y}$, e.g., $\mathbf{v}_j = \mathbf{y} + \mathbf{n}_j$ where $\mathbf{n}_j$ is a random noise. In this case, system filters signals $\mathbf{v}_0, \ldots, \mathbf{v}_p$.

Further, for $i = 0, 1, \ldots, p$, let $G_i \in R^{m \times r_i}$, $H_i \in R^{r_i \times q_i}$ where r_i is given, $0 < r_i < r$ and

$$r = r_0 + \ldots + r_p. \tag{2.4}$$

Here, r is a positive integer such that $r \leq \min\{m, n\}$. It is convenient to set $Q_0 = I$ and $\mathbf{z}_0 = \mathbf{v}_0 = \mathbf{y}$.

For a given reduction ratio

$$c = r/\min\{m, n\},$$

we consider a system model given by operator $\mathcal{T}_p : L^2(\Omega, \mathbf{R}^{q_0}) \times \ldots \times L^2(\Omega, \mathbf{R}^{q_p}) \to L^2(\Omega, \mathbf{R}^m)$ such that

$$\mathcal{T}_p(\mathbf{v}_0, \ldots \mathbf{v}_p) = G_0 H_{00} + \ldots + G_p H_p \mathbf{z}_p, \tag{2.5}$$

where $G_j : L^2(\Omega, \mathbf{R}^{r_j}) \to L^2(\Omega, \mathbf{R}^m)$ and $H_j : L^2(\Omega, \mathbf{R}^{q_j}) \to L^2(\Omega, \mathbf{R}^{r_j})$, for $j = 0, 1, \ldots, p$, are linear operators (i.e. G_j and H_j are represented by $m \times r_j$ and $r_j \times q_j$ matrices, respectively. Recall, we use the same symbol to define a matrix and the associated liner operator).

Importantly, operators $H_0, \ldots, H_p$ imply the dimensionality reduction of vectors $\mathbf{v}_0, \ldots, \mathbf{v}_p$ because $H_{ii} \in L^2(\Omega, \mathbf{R}^{r_i})$ where $0 < r_i < r \leq \min\{m, n\}$, for $i = 0, \ldots, p$.

We call p the degree of model $\mathcal{T}_p$. It is shown below that $\mathcal{T}_p$ approximates the system of interest $\mathcal{F} : L^2(\Omega, \mathbf{R}^n) \to L^2(\Omega, \mathbf{R}^m)$ with the accuracy represented by Theorems 8-6 below.

2.3. Statement of the problem

Let $\mathcal{F} : L^2(\Omega, \mathbf{R}^n) \to L^2(\Omega, \mathbf{R}^m)$ be a continuous map. We consider the problem as follows: Given $\mathbf{x}, \mathbf{y}$, $\mathbf{v}_1, \ldots, \mathbf{v}_p$ and $r_0, \ldots, r_p$, find matrices $G_0, H_0, \ldots, G_p, H_p$ and transformations $Q_0, \ldots, Q_p$ that solve

$$\min_{G_0, H_0, \ldots, G_p, H_p} \left\| \mathcal{F}(\mathbf{y}) - \sum_{j=0}^{p} G_j H_j \mathbf{z}_j \right\|_{\Omega}^2 \tag{2.6}$$

subject to

$$G_j \in R^{m \times r_j} \text{ and } H_j \in R^{r_j \times q_j}, \tag{2.7}$$

and

$$E_{z_i z_j} = \mathbf{O}, \quad \text{for } i \neq j, \tag{2.8}$$

where $i, j = 0, \ldots, p$ and $\mathbf{O}$ denotes the zero matrix (and the zero vector). Here, $\mathbf{z}_j$, for $j = 0, \ldots, p$, is represented by (2.3). We note that, for $F_j = G_j H_j$, the constraint in (2.7) can also be written as

$$rank\ F_j = r_j.$$

It will be shown in Section 3 below that the solution of problem (2.6) - (2.8) is determined under a special condition imposed on vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$. The condition is provided by Definition 3.

An assumption used in the methods of transformations of random signals is that the covariance matrices E_{xy} and E_{yy} are known (see, for example, [36, 15, 40, 65, 82]). Here, we adopt this assumption. As mentioned, in particular, in [46, 47, 2, 85, 39, 90], *a priori* knowledge of the covariances can come either from specific data models, or, after sample estimation during a training phase. Examples 1 and 3 in Sections 3.3 and 3.4 below illustrate this observation.

3. Solution of problem (2.6)-(2.8)

Definition 1. *Random vectors $\mathbf{z}_0, \ldots, \mathbf{z}_p$ are called* pairwise uncorrelated *if the condition in (2.8) holds for any pair of vectors $\mathbf{z}_i$ and $\mathbf{z}_j$, for $i \neq j$, where $i, j = 0, \ldots, p$. Two vectors $\mathbf{z}_i$ and $\mathbf{z}_j$ belonging to the set of the pairwise uncorrelated vectors are called* uncorrelated.

For $j = 0, \ldots, p$, let $\mathcal{N}(M^{(j)})$ be a null space of matrix $M^{(j)} \in R^{q_j \times q_j}$.

Definition 2. *Random vectors* $\mathbf{v}_0, \ldots, \mathbf{v}_p$ *are called* jointly independent *if*

$$M^{(0)}\mathbf{v}_0(\omega) + \ldots + M^{(p)}\mathbf{v}_p(\omega) = \mathbf{O},$$

almost everywhere in Ω*, only if* $\mathbf{v}_j(\omega) \in \mathcal{N}(M^{(j)})$*, for* $j = 0, \ldots, p$.

Definition 3. *Random vector* $\mathbf{v}_j$*, for* $j = 0, \ldots, p$*, is called the* well-defined injection *if*

$$\Gamma_{z_j} = E_{xz_j} E^{\dagger}_{z_j z_j} E_{z_j x} \neq \mathbf{O}, \tag{3.1}$$

where $\mathbf{z_j}$ *is defined by (2.3). Otherwise, injection* $\mathbf{v}_j$ *is called ill-defined.*

An explanation for introducing Definition 3 is provided by Remark 1 at the end of Section 3.2 below.

3.1. Determination of pairwise uncorrelated vectors

Theorem 1. *Let random vectors* $\mathbf{v}_0, \ldots, \mathbf{v}_p$ *be jointly independent. Then they are transformed to the pairwise uncorrelated vectors* $\mathbf{z}_0, \ldots, \mathbf{z}_p$ *by transformations* $Q_0, \ldots, Q_p$ *as follows:*

$$\mathbf{z}_0 = Q_0(\mathbf{v}_0) = \mathbf{v}_0 \quad \textit{and, for } j = 1, \ldots, p, \tag{3.2}$$

$$\mathbf{z}_j = Q_j(\mathbf{v}_j, Z_{j-1}) = \mathbf{v}_j - \sum_{k=0}^{j-1} E_{v_j z_k} E^{\dagger}_{z_k z_k} \mathbf{z}_k. \tag{3.3}$$

Proof. Suppose that the condition in (2.8) holds for $\mathbf{z}_0, \ldots, \mathbf{z}_{i-1}$. Then, for $\ell = 0, \ldots, i-1$,

$$\begin{aligned} E_{z_i z_\ell} &= E[(\mathbf{v}_i - \sum_{l=0}^{i-1} E_{v_i z_l} E^{\dagger}_{z_l z_l} \mathbf{z}_l)\mathbf{z}_\ell^T] \\ &= E_{v_i z_\ell} - \sum_{l=0}^{i-1} E_{v_i z_l} E^{\dagger}_{z_l z_l} E_{z_l z_\ell} \\ &= E_{v_i z_\ell} - E_{v_i z_\ell} E^{\dagger}_{z_\ell z_\ell} E_{z_\ell z_\ell} = \mathbf{O}. \end{aligned} \tag{3.4}$$

The latter is true because by Lemma 1 in [80],

$$E_{v_i z_\ell} E^{\dagger}_{z_\ell z_\ell} E_{z_\ell z_\ell} = E_{v_i z_\ell}.$$

Thus, by induction, (2.8) holds for any $i = 0, \ldots, p$. □

3.2. Determination of matrices $G_0, H_0, \ldots, G_p, H_p$ that solve problem (2.6)

First, recall the definition of a truncated SVD. Let the SVD of matrix $A \in R^{m\times n}$ be given by

$$A = U_A \Sigma_A V_A^T,$$

where $U_A = [u_1\ u_2\ \ldots u_m] \in R^{m\times m}, V_A = [v_1\ v_2\ \ldots v_n] \in R^{n\times n}$ are unitary matrices, and $\Sigma_A = (\sigma_1(A),\ \ldots,\ \sigma_{\min(m,n)}(A)) \in R^{m\times n}$ is a generalized diagonal matrix, with the singular values $\sigma_1(A) \geq \sigma_2(A) \geq \ldots \geq 0$ on the main diagonal. For $k < m$, $j < n$ and $\ell < \min(m,n)$, we denote $U_{A,k} = [u_1\ u_2\ \ldots u_k]$, $V_{A,j} = [v_1\ v_2\ \ldots v_j]$, $\Sigma_{A,\ell} = (\sigma_1(A), \ldots, \sigma_\ell(A))$, and write

$$\Pi_{A,L} = \sum_{k=1}^{rank\ (A)} u_k u_k^T \text{ and } \Pi_{A,R} = \sum_{j=1}^{rank\ (A)} v_j v_j^T.$$

For $r = 1, \ldots, rank\ (A)$,

$$[A]_r = \sum_{i=1}^{r} \sigma_i(A) u_i v_i^T \in R^{m\times n}, \tag{3.5}$$

is the truncated SVD of A. For $r \geq rank\ (A)$ we write $[A]_r = A\ (= A_{rank\ (A)})$.

Theorem 2. *Let $\mathbf{v}_0, \ldots, \mathbf{v}_p$ be well-defined injections and vectors $\mathbf{z}_0, \ldots, \mathbf{z}_p$ be pairwise uncorrelated. Then the minimal Frobenius norm solution to the problem in (2.6) is given, for $j = 0, \ldots, p$, by*

$$G_j = U_{\Gamma_{z_j}, r_j} \text{ and } H_j = U^T_{\Gamma_{z_j}, r_j} E_{xz_j} E^{\dagger}_{z_j z_j}, \tag{3.6}$$

Proof. For $j = 0, \ldots, p$, let $S_j = G_j H_j$, and let $S = [S_0, \ldots, S_p]$ and $w = [\mathbf{z}_0^T, \ldots, \mathbf{z}_p^T]^T$. Then

$$\|\mathcal{F}(\mathbf{y}) - \sum_{j=0}^{p} S_{jj}\|^2_{\Omega} = tr\Big\{(E_{xx} - E_{xw}S^T - SE_{wx} + SE_{ww}S^T)\Big\}, \tag{3.7}$$

where $E_{xw} = [E_{xz_0}, \ldots, E_{xz_p}]$ and by Theorem 1, $E_{ww} = diag[E_{z_0 z_0}, \ldots, E_{z_p z_p}]$ is block-diagonal matrix. Thus,

$$SE_{ww}S^T = S_0 E_{z_0 z_0} S_0^T + \ldots + S_p E_{z_p z_p} S_p^T,$$

$$SE_{wx} = S_0 E_{z_0 x} + \ldots + S_p E_{z_p x}.$$

Therefore, (3.7) implies

$$(3.8)\, \|\mathcal{F}(\mathbf{y}) - \sum_{j=0}^{p} G_j H_{jj}\|_{\Omega}^2 = \sum_{j=0}^{p} \|\mathcal{F}(\mathbf{y}) - G_j H_j \mathbf{z_j}\|_{\Omega}^2 - tr\{pE_{xx}\}.$$

Let us denote by $\mathbf{R}_r^{m\times n}$ the set of all $m \times n$ matrices of rank at most r. Then on the basis of [83, 30, 52, 12, 87], the minimal Frobenius norm solution to the problem

$$(3.9) \qquad \min_{G_j H_j \in \mathbf{R}_{r_j}^{m\times n}} \|\mathcal{F}(\mathbf{y}) - G_j H_j \mathbf{z}_j\|_{\Omega}^2,$$

for $j = 0, \ldots, p$, is given by

$$(3.10) \qquad G_j H_j = U_{\Gamma_{z_j}, r_j} U_{\Gamma_{z_j}, r_j}^T E_{x z_j} E_{z_j z_j}^{\dagger}$$

Therefore, (3.6) follows from (3.10). □

Remark 1. *Definition 3 of the well-defined injections is motivated by the following observation. It follows from (3.6) that if, for all $j = 0, \ldots, p$, vector $\mathbf{v}_j$ is such that $\Gamma_{z_j} = \mathbf{O}$, then $G_j = \mathbf{O}$ and $H_j = \mathbf{O}$. In other words, then approximating operator $\mathcal{T}_p = \mathbf{O}$.*

Therefore, in Theorem 2 above and in the Theorems below, vectors $\mathbf{v}_0, \ldots, \mathbf{v}_p$ are assumed well-defined.

3.3. Empirical determination of well-defined injections

First, we wish to show that the proposed system model $\mathcal{T}_p$ is also applicable under the conventional assumption used, for example, in [15, 36, 65, 83, 35, 70], when only matrices E_{xx}, E_{xy} and E_{yy} are assumed to be givenFootnoteIn particular, if $\mathbf{y} = A\mathbf{x} + \boldsymbol{\xi}$ where A is a known matrix and random vector $\boldsymbol{\xi}$ is independent on $\mathbf{x}$, then it is assumed that only E_{xx} and $E_{\mathbf{x}i,\mathbf{x}i}$ are given., and matrices E_{xv_j}, E_{yv_j}, $E_{v_\ell v_j}$ and associated joint probability functions are unknown. In this case, E_{xv_j}, E_{yv_j} and $E_{v_\ell v_j}$ should be represented in terms of E_{xx}, E_{xy} and E_{yy}.

We use the Stein's lemma [76] to show that it can be done indeed and, therefore, suppose that, for $q_j \geq n$, vector $\mathbf{y} = [\mathbf{y}_{(1)}, \ldots, \mathbf{y}_{(n)}]^T$ is extended to vector $\widetilde{\mathbf{y}}^{(j)}$ with q_j components, $\widetilde{\mathbf{y}}^{(j)} = [\widetilde{\mathbf{y}}_1^{(j)}, \ldots, \widetilde{\mathbf{y}}_{q_j}^{(j)}]^T$, where s_k entries of $\widetilde{\mathbf{y}}^{(j)}$ coincide with $\mathbf{y}_{(k)}$, for $k = 1, \ldots, n$, and $q_j =$

$s_1 + \ldots + s_n$FootnoteThere are $q_j!/s_1!\ldots s_n!$ such orderings.. In fact, $s_1, \ldots, s_n$ depend on j but we omit subscript j for the notation simplicity. (For example, if $\mathbf{y} = [\mathbf{y}_{(1)}, \mathbf{y}_{(2)}]^T$ then one of such extensions is $\widetilde{\mathbf{y}}^{(j)} = [\widetilde{\mathbf{y}}_1^{(j)}, \widetilde{\mathbf{y}}_2^{(j)}, \widetilde{\mathbf{y}}_3^{(j)}]^T = [\mathbf{y}_1^{(j)}, \mathbf{y}_2^{(j)}, \mathbf{y}_1^{(j)}]^T$ where $s_1 = 2$ and $s_2 = 1$.)

In general, injection $\mathbf{v}_j$, for $j = 1, \ldots, p$, is defined by a nonlinear transformation f_j so that $\mathbf{v}_j = f_j(\mathbf{y}) \in L^2(\Omega, \mathbf{R}^{q_j})$ and $q_j \geq n$. Therefore, $\mathbf{v}_j = f_j(\mathbf{y})$ can be represented as $\mathbf{v}_j = \varphi_j[\phi_j(\mathbf{y})] = \varphi_j[\widetilde{\mathbf{y}}^{(j)}]$ where $\widetilde{\mathbf{y}}^{(j)} = \phi_j(\mathbf{y})$, $\mathbf{v}_j = [\mathbf{v}_1^{(j)}, \ldots, \mathbf{v}_{q_j}^{(j)}]^T$ and $\mathbf{v}_k^{(j)} = \varphi_{jk}(\widetilde{\mathbf{y}}_k^{(j)})$, for $k = 1, \ldots, q_j$, i.e., $[\mathbf{v}_1^{(j)}, \ldots, \mathbf{v}_{q_j}^{(j)}]^T = [\varphi_{jk}(\widetilde{\mathbf{y}}_1^{(j)}), \ldots, \varphi_{jk}(\widetilde{\mathbf{y}}_{q_j}^{(j)})]^T$.

By the conditions of the Stein's lemma [76], we also suppose that

(A1) φ_j is a smooth function,

(A2) $E[|\varphi_j^{(n)}(\widetilde{\mathbf{y}}_j)|] < \infty$,

(A3) $E[\varphi_j'(\widetilde{\mathbf{y}}_j)] \neq 0$,

(A4) sequence $\{a_n\}_{n=0}^{\infty}$ where $a_n = E[\varphi_i^{(n)}(\widetilde{\mathbf{x}}_i)] \times E[\varphi_j^{(n)}(\widetilde{\mathbf{x}}_j)]$ is monotone and bounded, for all $i, j = 1, ..., q$, and

(A5) $\mathbf{y}_i$ and $\varphi_j(\widetilde{\mathbf{y}}_k^{(j)})$ are jointly normally distributed, for all $i, j = 1, ..., q$.

Let us write $E_{y\widetilde{y}^{(j)}} = \{E_{y_i\widetilde{y}_k^{(j)}}\}_{i,k=1}^{n,q_j}$. Then by the Stein's lemma,

$$E_{y_i v_k^{(j)}} = E_{y_i \varphi_{jk}(\widetilde{\mathbf{y}}_k^{(j)})} = E_{y_i \widetilde{y}_k^{(j)}} E[\varphi_{jk}'(\widetilde{\mathbf{y}}_k^{(j)})].$$

Therefore,

$$E_{yv_j} = E_{y\widetilde{y}^{(j)}} D_j, \tag{3.11}$$

where

$$D_j = \mathrm{diag}(E[\varphi_{j1}'(\widetilde{\mathbf{y}}_1^{(j)})], \ldots, E[\varphi_{jq_j}'(\widetilde{\mathbf{y}}_{q_j}^{(j)})]) \neq \mathbf{O}.$$

Similarly,

$$E_{xv_j} = E_{x\widetilde{y}^{(j)}} D_j. \tag{3.12}$$

Thus, $E_{yv_j} \neq \mathbf{O}$ and $E_{xv_j} \neq \mathbf{O}$. Now, to obtain $E_{v_\ell v_j}$, we need to compute each entry $E_{v_i^{(\ell)} v_k^{(j)}} = E_{\varphi_{\ell i}(\widetilde{y}_i^{(\ell)})\, \varphi_{jk}(\widetilde{y}_k^{(j)})}$, for $i, k = 1, 2, ..., q_j$. By the results obtained in Section 3 of [59],

$$E_{\varphi_{\ell i}(\widetilde{y}_i^{(j)})\, \varphi_{jk}(\widetilde{y}_k^{(j)})} = \sum_{n=1}^{\infty} \frac{\left(E_{\widetilde{y}_i^{(\ell)} \widetilde{y}_k^{(j)}}\right)^n}{n!} E[\varphi_{\ell i}^{(n)}(\widetilde{\mathbf{y}}_i^{(\ell)})] E[\varphi_{jk}^{(n)}(\widetilde{\mathbf{y}}_k^{(j)})]. \tag{3.13}$$

On the basis of assumption (A4) and the Abel's test (see, e.g., Corollary 5.58 in [64]), the series in (3.13) converges.

In (3.11), (3.12) and (3.13), entries of matrices $E_{y\widetilde{y}^{(j)}}$, $E_{x\widetilde{y}^{(j)}}$ and $E_{\widetilde{y}_i^{(j)}\widetilde{y}_k^{(j)}}$ are the same as entries of matrices E_{yy}, E_{xy} and E_{yy}, respectively. Therefore, E_{yv_j}, E_{xv_j} and $E_{v_\ell v_k}$ are determined, in fact, in terms of E_{xy} and E_{yy}, and expectations $E[\varphi'_{\ell i}(\widetilde{\mathbf{y}}_i^{(\ell)})]$ and $E[\varphi_{jk}^{(n)}(\widetilde{\mathbf{y}}_k^{(j)})]$, for $i,k=1,\ldots,q_j$.

The following Example 1 illustrates the above derivations. Injections $\mathbf{v}_1,\ldots,\mathbf{v}_p$ defined in Example 1, are well-defined in the sense of condition (3.1) checked up numerically.

Example 1. *Random vectors of type* $[e^{\widetilde{\mathbf{y}}_1},\ldots,e^{\widetilde{\mathbf{y}}_q}]^T$ *have been exploited, in particular, in [43]. In this regard, let us consider* $\mathbf{v}_k^{(j)}=\varphi_{jk}(\widetilde{\mathbf{y}}_k^{(j)})=e^{\widetilde{\mathbf{y}}_k^{(j)}}$, *for* $j=1,..,p$ *and* $k=1,\ldots,q_j$, *and suppose the above conditions (A1)-(A5) hold. Then* $E[\varphi'_{jk}(\widetilde{\mathbf{y}}_k^{(j)}]=E[\varphi_{jk}(\widetilde{\mathbf{y}}_k^{(j)}]=E[e^{\widetilde{\mathbf{y}}_k^{(j)}}]=\sqrt{e}$. *Therefore,* $D_j=\sqrt{e}I$ *and then* $E_{yv_j}=\sqrt{e}E_{y\widetilde{y}^{(j)}}$ *and* $E_{xv_j}=\sqrt{e}E_{x\widetilde{y}^{(j)}}$. *On the other hand,* $E[\varphi_{ji}^{(n)}(\widetilde{\mathbf{y}}_i^{(j)})]=E[\varphi_{ji}(\widetilde{\mathbf{y}}_i^{(j)})]=\sqrt{e}$. *Thus, (3.13) implies*

$$E_{v_i^{(\ell)}v_k^{(j)}}=E_{\varphi_{\ell i}(\widetilde{y}_i^{(\ell)})\,\varphi_{jk}(\widetilde{y}_k^{(j)})}=e\sum_{n=1}^{\infty}\frac{\left(E_{\widetilde{y}_i^{(\ell)}\widetilde{y}_k^{(j)}}\right)^n}{n!}=e^{E_{\widetilde{y}_i^{(\ell)}\widetilde{y}_k^{(j)}}+1}-e.$$

Therefore, $E_{v_\ell v_k}=eM_{\ell k}$ *where* $M_{\ell k}=\{e^{E_{\widetilde{y}_i^{(\ell)}\widetilde{y}_k^{(j)}}}-1\}_{i,k=1}^{q_j}$. *Then, in particular, for* $j=1$, $E_{xz_1}=E_{xv_1}-E_{xy}E_{yy}^{\dagger}E_{yv_1}=\sqrt{e}(E_{x\widetilde{y}^{(1)}}-E_{xy}E_{yy}^{\dagger}E_{y\widetilde{y}^{(1)}})$ *and* $E_{z_1z_1}=e(M_{11}-E_{\widetilde{y}^{(1)}y}E_{yy}^{\dagger}E_{y\widetilde{y}^{(1)}})$, *where* $\widetilde{\mathbf{y}}^{(1)}\neq\mathbf{y}$. *That is, condition (3.1) is represented in terms of entries of matrices* E_{xy} *and* E_{yy}, *as supposed, and can be verified numerically.*

Remark 2. *In our simulations in Examples 2, 3 and 4 below, empirical injections* $\mathbf{v}_1,\ldots,\mathbf{v}_p$ *are chosen 'similar' to vector* $\mathbf{x}$. *That is, if by the assumption,* $\mathbf{x}$ *is, e.g., Gaussian then* $\mathbf{v}_1,\ldots,\mathbf{v}_p$ *are Gaussian as well. This is another way to define* $\mathbf{v}_1,\ldots,\mathbf{v}_p$.

3.4. Error analysis associated with solution of problem (2.6)-(2.8)

Here, in Theorem 3, we obtain the constructive representation of the error associated with the solution of problem (2.6)-(2.8). We also prove that the error can be improved by the increase in the dimensions of injections

$\mathbf{v}_1, \ldots, \mathbf{v}_p$ (Theorem 4) and degree of approximating operator $\mathcal{T}_p$ (Theorems 3, 4, 5).

Let us write

$$\varepsilon_{GH}^{(p)} = \min_{G_0,H_0,\ldots,G_p,H_p} \|\mathcal{F}(\mathbf{y}) - \sum_{j=0}^{p} G_j H_{jj}\|_\Omega^2.$$

Theorem 3. *For $j = 0, \ldots, p$, let $A_j = E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger$, rank $(A_j) = s_j$ and $s_j \geq r_j + 1$. For $k = 1, \ldots, s_j$, let σ_k be a singular value of A_j. Then the error associated with the solution of problem (2.6)-(2.8) is represented by*

$$\varepsilon_{GH}^{(p)} = \mathrm{tr}\{E_{xx}\} - \sum_{j=0}^{p}\sum_{k=1}^{r_j} \sigma_k^2(A_j) \tag{3.14}$$

In particular, the error decreases as p increases.

Proof. For G_jH_j determined by (3.10), we have

$$\begin{aligned}
\|\mathcal{F}(\mathbf{y}) - G_jH_j\mathbf{z_j}\|_\Omega^2 &= tr\{E_{xx} - E_{xz_j}(G_jH_j)^T - G_jH_jE_{z_jx} + G_jH_jE_{z_jz_j}(G_jH_j)^T\} \\
&= \|E_{xx}^{1/2}\|^2 - \|E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger\|^2 + \|(G_jH_j - E_{xz_j}E_{z_jz_j}^\dagger)E_{z_jz_j}^{1/2}\|^2 \\
&= \|E_{xx}^{1/2}\|^2 - \|E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger\|^2 + \|G_jH_jE_{z_jz_j}^{1/2} - E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger\|^2 \qquad (3.15)
\end{aligned}$$

because

$$E_{z_jz_j}^\dagger E_{z_jz_j}^{1/2} = (E_{z_jz_j}^{1/2})^\dagger$$

and

$$E_{xz_j}E_{z_jz_j}^\dagger E_{z_jz_j} = E_{xz_j} \tag{3.16}$$

(see [80]). Further, it follows from (3.6) that in the notation introduced in (3.5), matrix G_jH_j is represented as $G_jH_j = [E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger]_{r_j}(E_{z_jz_j}^{1/2})^\dagger$. Therefore,

$$\begin{aligned}
\|G_jH_jE_{z_jz_j}^{1/2} - E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger\|^2 &= \|[E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger]_{r_j}(E_{z_jz_j}^{1/2})^\dagger E_{z_jz_j}^{1/2} - E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger\|^2 \\
&= \|[E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger]_{r_j} - E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger\|^2 \\
&= \|[A_j]_{r_j} - A_j\|^2 \\
&= \sum_{k=r_j+1}^{s_j} \sigma_k^2(A_j). \qquad (3.17)
\end{aligned}$$

because $[E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger]_{r_j}(E_{z_jz_j}^{1/2})^\dagger E_{z_jz_j}^{1/2} = [E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger]_{r_j}$. Further, since

$$\|E_{xz_j}(E_{z_jz_j}^{1/2})^\dagger\|^2 = \|A_j\|^2 = \sum_{k=1}^{s_j}\sigma_k^2(A_j), \tag{3.18}$$

then (3.8), (3.15), (3.17) and (3.18) imply (3.14). In particular, it follows from (3.14) that the error decreases as degree p increases. □

Remark 3. *According to Remark 2, in Examples 2, 3 and 4 that follow, empirical injections $\mathbf{v}_1,\dots,\mathbf{v}_p$ are chosen 'similar' to vector $\mathbf{x}$. They are well-defined in the sense of condition (3.1) verified numerically.*

Further, we wish to show the utility of injections $\mathbf{v}_1,\dots,\mathbf{v}_p$ in the sense of the diminution of the associated error by increasing dimensions of $\mathbf{v}_1,\dots,\mathbf{v}_p$. To this end, we write

$$A_j = \{a_{ki(j)}\}_{k,i=1}^{m,q_j} \text{ and } A_j - [A_j]_{r_j} = \{b_{ki(j)}\}_{k,i=1}^{m,q_j}$$

where $a_{ki(j)}$ and $a_{ki(j)}$ are entries of matrices A_j and $A_j - [A_j]_{r_j}$, respectively. Let us also denote

$$\gamma_{k,(j)} = \sum_{i=1}^{m}\left(a_{ki(j)}^2 - b_{ki(j)}^2\right),$$

$$\gamma_{(j)} = \max\{\gamma_{1,(j)},\dots,\gamma_{q_j,(j)}\},$$

$$\gamma = \max_{j=1,\dots,p}\gamma_{(j)},$$

$\alpha_0 = \operatorname{tr}\{E_{xx} - \sum_{j=0}^{p} A_jA_j^T\}$, and $q = q_1 + \dots, +q_p$. We wish to show the utility of injections $\mathbf{v}_1,\dots,\mathbf{v}_p$ in the sense of the diminution of the associated error by increasing dimensions of $\mathbf{v}_1,\dots,\mathbf{v}_p$.

Theorem 4. *Let $\mathbf{v}_1,\dots,\mathbf{v}_p$ be well-defined injections and let matrices $G_0, H_0, \dots, G_p, H_p$ be defined by Theorem 2. Then the associated error decreases as the sum q of dimensions of injections $\mathbf{v}_1,\dots,\mathbf{v}_p$ increases. In particular, there is $\beta \in (0,\gamma]$ such that, given $\alpha \geq \alpha_0$, then*

$$\alpha_0 \leq \varepsilon_{GH}^{(p)} \leq \alpha \quad \textit{iff} \quad q \geq \frac{\operatorname{tr}\{E_{xx}\} - \sum_{k=1}^{r_0}\sigma_k^2(A_0)}{\beta}. \tag{3.19}$$

Proof. It follows from (3.8), (3.14), (3.15) and (3.17) that

$$\begin{aligned}
\varepsilon_{GH}^{(p)} &= \|\mathcal{F}(\mathbf{y}) - G_j H_0 \mathbf{z_0}\|_\Omega^2 + \sum_{j=1}^{p} \|\mathcal{F}(\mathbf{y}) - G_j H_j \mathbf{z_j}\|_\Omega^2 - tr\{pE_{xx}\} \\
&= tr\{E_{xx}\} - \sum_{k=1}^{r_0} \sigma_k^2(A_0) + \sum_{j=1}^{p} \|\mathcal{F}(\mathbf{y}) - G_j H_j \mathbf{z_j}\|_\Omega^2 - tr\{pE_{xx}\} \qquad (3.20)
\end{aligned}$$

where

$$\begin{aligned}
\|\mathcal{F}(\mathbf{y}) - G_j H_j \mathbf{z_j}\|_\Omega^2 &= tr\{E_{xx}\} - \left(\|A_j\|^2 - \left\| [A_j]_{r_j} - A_j \right\|^2 \right) \\
&= tr\{E_{xx}\} - \sum_{k=1}^{q_j} \sum_{i=1}^{m} \left(a_{ki(j)}^2 - b_{ki(j)}^2 \right).
\end{aligned}$$

Therefore,

$$\varepsilon_{GH}^{(p)} = tr\{E_{xx}\} - \sum_{k=1}^{r_0} \sigma_k^2(A_0) - \sum_{j=1}^{p} \sum_{k=1}^{q_j} \sum_{i=1}^{m} \left(a_{ki(j)}^2 - b_{ki(j)}^2 \right). \qquad (3.21)$$

Here, $\sum_{k=1}^{q_j} \sum_{i=1}^{m} \left(a_{ki(j)}^2 - b_{ki(j)}^2 \right) > 0$ since by (3.17),

$$\|A_j\|^2 - \left\| [A_j]_{r_j} - A_j \right\|^2 = \sum_{k=1}^{r_j} \sigma_k^2(A_j) > 0. \qquad (3.22)$$

Thus, (3.20) - (3.22) imply that $\varepsilon_{GH}^{(p)}$ decreases as q_j increases, for $j = 1, \ldots, p$, and p increases. Further, (3.14) implies

$$\varepsilon_{GH}^{(p)} \geq tr\{E_{xx}\} - \sum_{j=0}^{p} \sum_{k=1}^{s_j} \sigma_k^2(A_j) = tr\left\{ E_{xx} - \sum_{j=0}^{p} A_j A_j^T \right\} = \alpha_0. \qquad (3.23)$$

Since

$$0 < \sum_{j=1}^{p} \sum_{k=1}^{q_j} \sum_{i=1}^{m} \left(a_{ki(j)}^2 - b_{ki(j)}^2 \right) \leq \gamma \sum_{j=1}^{p} q_j = \gamma q,$$

then

$$\sum_{j=1}^{p} \sum_{k=1}^{q_j} \sum_{i=1}^{m} \left(a_{ki(j)}^2 - b_{ki(j)}^2 \right) = q\beta. \qquad (3.24)$$

Therefore, (3.21), (3.23) and (3.24) imply

$$\alpha_0 \leq \varepsilon_{GH}^{(p)} = tr\{E_{xx}\} - \sum_{k=1}^{r_0} \sigma_k^2(A_0) - q\beta.$$

thus, if $\varepsilon_{GH}^{(p)} \leq \alpha$ then the right inequality in (3.19) is true. Conversely, if the latter is true then $\varepsilon_{GH}^{(p)} \leq \alpha$. □

Remark 4. *An empirical explanation of Theorem 4 is that the increase in q implies the increase in the dimensions of matrices $H_1, \ldots, H_p$ in (2.5) and (3.6). Hence, it implies the increase in the number of parameters to optimize. As a result, for the given r, the accuracy associated with the system model $\mathcal{T}_p$ improves. Further, it follows from (3.19) that, as q increases, $\varepsilon_{GH}^{(p)}$ tends to α_0 which is the error associated with the full rank model $\mathcal{T}_h$ (see (3.32)). The error is given by (3.36).*

Remark 5. *By Theorem 3, the error associated with solution of problem (2.6) decreases as degree p of the system model $\mathcal{T}_p$ increases. At the same time, the increase in degree p of model $\mathcal{T}_p$ may involve an increase in parameter r given by (2.4). However, by a condition of some applied problem in hand, r must be fixed. In the following Theorem 5, under the condition of fixed r, the case of decreasing the error as the degree p of the system model increases is detailed.*

Theorem 5. *Let r and r_j, for $j = 0, \ldots, p$, be given. Let g be a nonnegative integer such that $g < p$ and let $\ell_g = r_g + r_{g+1} + \ldots + r_p$. If*

$$\sum_{k=r_g+1}^{r_{g+1}+\ldots+r_p} \sigma_k^2(A_g) < \sum_{j=g+1}^{p} \sum_{k=1}^{r_j} \sigma_k^2(A_j), \tag{3.25}$$

where

$$\sum_{j=g+1}^{p} \sum_{k=1}^{r_j} \sigma_k^2(A_j) = \sum_{j=g+1}^{p} \sum_{k=1}^{q_j} \sum_{i=1}^{m} \left(a_{ki(j)}^2 - b_{ki(j)}^2 \right),$$

then

$$\varepsilon_{GH}^{(p)} < \varepsilon_{GH}^{(g)}, \tag{3.26}$$

i.e., for the same r, the error associated with the higher degree model p is less than that associated with the lower degree model g.

Proof. We write $r = r_0 + \ldots + r_{g-1} + \ell_g$. Then

$$\begin{aligned} \varepsilon_{GH}^{(g)} &= tr\{E_{xx}\} - \sum_{j=0}^{g-1}\sum_{k=1}^{r_j}\sigma_k^2(A_j) - \sum_{k=1}^{\ell_g}\sigma_k^2(A_g) \\ (3.27) \quad &= tr\{E_{xx}\} - \sum_{j=0}^{g-1}\sum_{k=1}^{r_j}\sigma_k^2(A_j) - \sum_{k=1}^{r_g}\sigma_k^2(A_g) - \sum_{k=r_g+1}^{r_{g+1}+\ldots+r_p}\sigma_k^2(A_g). \end{aligned}$$

Thus, (3.14) and (3.27) imply (3.25) and (3.26). □

Remark 6. *The RHS in (3.25) increases as the dimension q_j of at least single injection $\mathbf{v}_j$, for $j = g+1, \ldots, p$, increases while the LHS does not depend on q_j. In other words, one can always find q_j, for $j = g+1, \ldots, p$, such that the inequality in (3.25) is true.*

Remark 7. *If $p = 1$ and $m = q_0 = q_1$, then proposed method is equivalent to the GBT2 [83]. For $m = q_0 = q_1$ and $r_0 = r_1 = m/4$, the computational load of the GBT2 is estimated as $57m^3$. The computational load associated with computation of G_j and H_j by (3.6) is represented in Table 3.1.*

Table 3.1: Estimate of computational load associated with computation of G_j and H_j.

Particular parts of (3.6)	**Number of flops**
1. $E^{\dagger}_{z_j z_j}$	$14q_j^3$
2. $E_{xz_j}E^{\dagger}_{z_j z_j}$	$2mq_j^2$
3. $\Gamma_{z_j} = E_{xz_j}E^{\dagger}_{z_j z_j}E_{z_j x}$	$2m^2q_j$
4. $U_{\Gamma_{z_j}}$ *from SVD of* Γ_{z_j}	$12m^3$
5. $H_j = U^T_{T_{z_j},r_j}E_{xz_j}E^{\dagger}_{z_j z_j}$	$2m^2q_j + m^2r_j$

It follows from Table 3.1 that the total computational load of the proposed method is given by

$$\sum_{j=0}^{p}(14q_j^3 + 2mq_j^2 + 4m^2q_j + m^2r_j + 12m^3).$$

In particular, if $m = q_0 = ... = q_p$ and $r_j = m/4$, for all $j = 0, 1, ..., p$, then the computational load of the proposed method is $32pm^3$. A lower error is provided by the proposed method at the expense of increased computational and/or storage costs what, we believe, is quite natural. In general, a better

Table 3.2: Numerical characterizations of models $\mathcal{T}_0(\mathbf{v}_0)$, $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$ and $\mathcal{T}_2(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$ in Case 1

Model	q_0	q_1	q_2	r_0	r_1	r_2	MSE
$\mathcal{T}_0(\mathbf{v}_0)$	100	N/A	N/A	50	N/A	N/A	8.30
$\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$	100	25	N/A	25	25	N/A	7.93
$\mathcal{T}_2(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$	100	25	500	17	17	16	7.03

Table 3.3: Numerical characterizations of models $\mathcal{T}_0(\mathbf{v}_0)$, $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$ and $\mathcal{T}_2(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$ in Case 2

Model	q_0	q_1	q_2	r_0	r_1	r_2	MSE
$\mathcal{T}_0(\mathbf{v}_0)$	100	N/A	N/A	50	N/A	N/A	8.30
$\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$	100	200	N/A	25	25	N/A	7.61
$\mathcal{T}_2(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$	100	200	500	17	17	16	6.28

result is normally achieved by some additional efforts. This observation is illustrated by the following example.

Example 2. *Here, we wish to numerically illustrate Theorems 3, 4 and 5. To this end, we assume that* $\mathbf{y} = \mathbf{x} + \boldsymbol{\xi}$ *where* $\mathbf{x} \in L^2(\Omega, R^m)$ *and* $\boldsymbol{\xi} \in L^2(\Omega, R^m)$ *are uniformly and normally distributed random vectors, respectively. Injections* $\mathbf{v}_1 \in L^2(\Omega, R^{q_1})$ *and* $\mathbf{v}_2 \in L^2(\Omega, R^{q_2})$ *are here chosen as uniformly distributed random vectors. Random vector* $\boldsymbol{\xi}$ *simulates noise and is uncorrelated with* $\mathbf{x}$ *and* $\mathbf{v_1}, \mathbf{v_2}$*. Covariance matrices* E_{xv_j}, $E_{v_iv_j}$, *for* $i, j = 0, 1, 2$, *are represented by* $E_{xv_j} = \frac{1}{s} X V_j^T$ *and* $E_{v_iv_j} = \frac{1}{s} V_i V_j^T$ *where* $X \in \mathbf{R}^{m \times s}$ *and* $V_j \in \mathbf{R}^{q_j \times s}$ *are sample matrices of* $\mathbf{x}$ *and* $\mathbf{v}_j$, *respectively, for* $j = 0, 1, 2$. *We choose* $m = 100$ *and* $r = 50$. *Then models* $\mathcal{T}_0(\mathbf{v}_0)$, $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$ *and* $\mathcal{T}_2(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$ *are specified by choosing particular values of* q_j *and* r_j, *for* $j = 0, 1, 2$, *as shown in Tables 1 and 2.*

In the Tables, Cases 1 and 2 for specific values of q_j *and* r_j, *for* $j = 0, 1, 2$, *are considered. The values of the associated mean square errors (MSE) are also included. In Figure 3.1 the MSE are illustrated diagrammatically. It follows from Tables 1,2 and Figure 3.1 that, for the same* r, *the error associated with the proposed system model decreases if degree* p *or the sum* q *of the injection dimensions increases.*

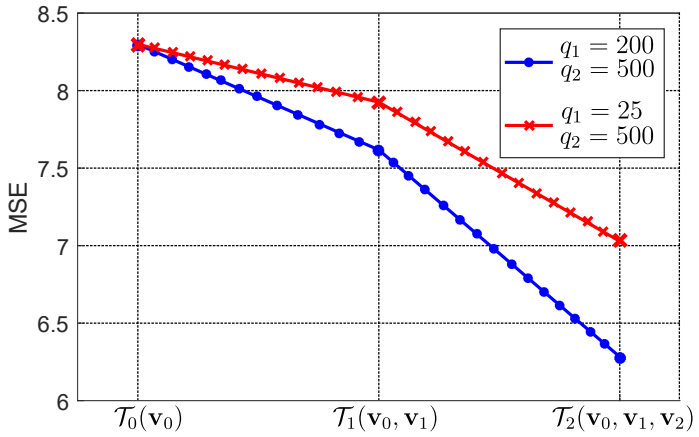


Figure 3.1: Example 2: Diagrams of the errors associated with $\mathcal{T}_0(\mathbf{v}_0)$, $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$ and $\mathcal{T}_2(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$.

Here, we wish to compare the error associated with the proposed technique (represented by $\varepsilon_{GH}^{(p)}$ in Theorem 3) with that of the method considered in [83] and abbreviated as GBT2. Recall that GBT2 is represented by $B(\mathbf{y}, g) = B_1(B_2\mathbf{y} + B_3 g)$ where $B : L^2(\Omega, \mathbf{R}^{q_0}) \times L^2(\Omega, \mathbf{R}^{q_1}) \to L^2(\Omega, \mathbf{R}^m)$, $g \in L^2(\Omega, \mathbf{R}^{q_1})$, and B_1, B_2, B_3 are linear operators defined by matrices $B_1 \in R^{m \times m}$ and $B_j \in R^{m \times q_{j-2}}$, for $j = 2, 3$. GBT2 is not assigned for a system modeling and for the problem in (2.6), (2.7) and (2.8). At the same time, in terms of the model degree introduced in Section 2.2, GBT2 can be interpreted as a 'second' degree model. Therefore, it is interesting to compare the associated errors. Let $\mathbf{u} = [\mathbf{y}^T g^T]^T \in L^2(\Omega, \mathbf{R}^{2n})$ and denote by ε_{GBT2} the error associated with GBT2.

Theorem 6. *Let* $A_u = E_{xu}(E_{uu}^{1/2})^{\dagger}$. *If*

$$\sum_{j=0}^{p} \sum_{k=1}^{r_j} \sigma_k^2(A_j) > \sum_{\ell=1}^{r} \sigma_\ell^2(A_u) \tag{3.28}$$

then

$$\varepsilon_{GH}^{(p)} < \varepsilon_{GBT2}. \tag{3.29}$$

Proof. By [83], $\varepsilon_{GBT2} = \mathrm{tr}\{E_{xx}\} - \sum_{\ell=1}^{r} \sigma_\ell^2(A_u)$. Then the latter and (3.14) involve the theorem validity. □

In other words, under the condition in (3.28), the error associated with the proposed technique is less than that of GBT2. Note that the condition in (3.28) is easy to satisfy by the increase in p, for example.

Remark 8. *The following Example 3 numerically illustrates the theoretical results obtained above and also demonstrates how the proposed technique is extended to the case of complex-valued random vectors.*

Example 3. *Let $\mathbf{x}_c \in L^2(\Omega, \mathbf{C}^m)$ and $\mathbf{y}_c \in L^2(\Omega, \mathbf{C}^m)$, i.e., let $\mathbf{x}_c$ and $\mathbf{y}_c$ be random vectors with complex values. Then $\mathbf{x}_c = \mathbf{x}_r + j\mathbf{x}_i$ and $\mathbf{y}_c = \mathbf{y}_r + j\mathbf{y}_i$ where $j = \sqrt{-1}$, and $\mathbf{x}_r, \mathbf{y}_r \in L^2(\Omega, \mathbf{R}^m)$ and $\mathbf{x}_i, \mathbf{y}_i \in L^2(\Omega, \mathbf{R}^m)$ are the real and imaginary parts of $\mathbf{x}_c$ and $\mathbf{y}_c$, respectively, represented by Gaussian random vectors with zero-mean and standard deviation one (abbreviated as $G(0,1)$). Let us consider the case when $\mathbf{y}_c$ is represented by*

$$\mathbf{y}_c = A_c\mathbf{x}_c + \boldsymbol{\xi}_c, \tag{3.30}$$

where $A_c = diag(e^{j\phi_1}, e^{j\phi_2}, ..., e^{j\phi_m}) \in \mathbf{C}^{m\times m}$, $j = \sqrt{-1}$, ϕ_i is a random number generated from Gaussian distribution with zero-mean and standard deviation one, $\boldsymbol{\xi}_c \in L^2(\Omega, \mathbf{C}^m)$ with $E_{\mathbf{x}i_c\mathbf{x}i_c} = \sigma^2 I$ and $\sigma \in R$. The random vectors of this type are considered in a number of applications such as those in [73, 49, 53, 33, 58].

As shown, in particular, in [61, 1], the problems concerning complex random vector $\mathbf{x}_c$ can always be treated in terms of the associated real random vector

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_r \\ \mathbf{x}_i \end{bmatrix} \in L^2(\Omega, \mathbf{R}^{2m}).$$

Let $\boldsymbol{\xi}_c = \boldsymbol{\xi}_r + j\boldsymbol{\xi}_i$ and $A_c = A_r + jA_i$. The complex-valued vector given by (3.30) holds if and only if the real-valued random vector $\mathbf{y} = A\mathbf{x} + \boldsymbol{\xi}$ holds, where

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_r \\ \mathbf{y}_i \end{bmatrix} \in L^2(\Omega, \mathbf{R}^{2m}), \quad \boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\xi}_r \\ \boldsymbol{\xi}_i \end{bmatrix} \in L^2(\Omega, \mathbf{R}^{2m}),$$

and

$$A = \begin{bmatrix} A_r & -A_i \\ A_i & A_r \end{bmatrix} \in \mathbf{R}^{2m\times 2m}.$$

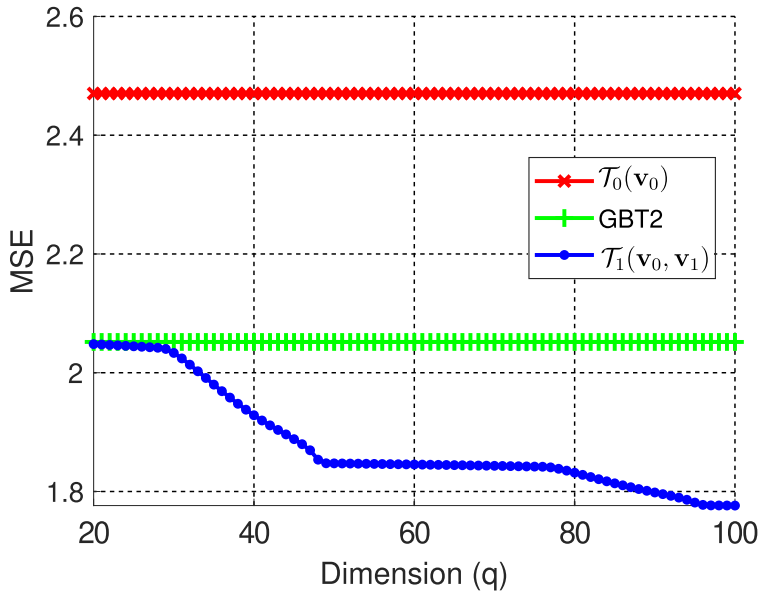
See, for instance, Appendix II in [61]. The above implies

$$E_{xx} = \begin{bmatrix} E_{x_r x_r} & E_{x_r x_i} \\ E_{x_i x_r} & E_{x_i x_i} \end{bmatrix} \in \mathbf{R}^{2m \times 2m}$$

and $E_{\mathbf{x}i\mathbf{x}i} = \sigma^2 I$. *Further, we denote* $E_{xx} = \{e_{k,\ell}\}_{k,\ell=1}^{2m}$. *In applications such as those associated with synthetic aperture radar interferometry (see, e.g., [53, 33]) matrix* E_{xx} *has a so-called triangular-time-Gaussian shape generated by*

$$e_{k,\ell} = \begin{cases} 1 - \dfrac{|k-\ell|}{m-1} b_k, & \text{for } |k-\ell| \le \dfrac{m-1}{b_k}, \\ 0, & \text{otherwise}, \end{cases} \tag{3.31}$$

where b_k *is a parameter (called in [53, 33] the normalized baseline parameter). Here,* E_{xx} *is considered in this form.*



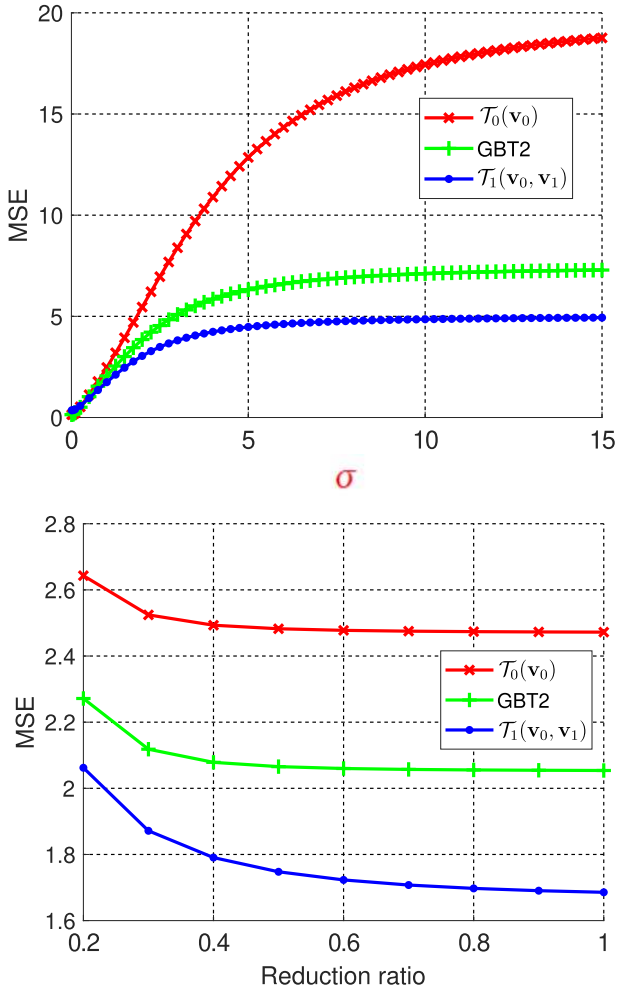


Figure 3.2: Example 3: Diagrams of the errors associated with $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$, $\mathcal{T}_0(\mathbf{v}_0)$ and GBT2 [83].

In this example, the injection $\mathbf{v}_1$ *is represented as complex random vector* $\mathbf{v}_{1c} = \mathbf{v}_{1r} + j\mathbf{v}_{1i} \in L^2(\Omega, \mathbf{C}^{q_1})$, *where* $\mathbf{v}_{1r} \in L^2(\Omega, \mathbf{R}^{q_1})$ *and* $\mathbf{v}_{1i} \in L^2(\Omega, \mathbf{R}^{q_1})$ *are the* $G(0,1)$. *It is assumed that* $\mathbf{v}_{1c}$ *is uncorrelated with* $\boldsymbol{\xi}_c$. *Further, we write*

$$\mathbf{v}_1 = \begin{bmatrix} \mathbf{v}_{1r} \\ \mathbf{v}_{1i} \end{bmatrix} \in L^2(\Omega, \mathbf{R}^{2q_1}), \qquad \mathbf{u} = \begin{bmatrix} \mathbf{x} \\ \mathbf{v}_1 \end{bmatrix} \in L^2(\Omega, \mathbf{R}^{2(m+q_1)}),$$

and choose $\mathbf{v}_1$ *such that covariance matrix*

$$E_{uu} = \begin{bmatrix} E_{xx} & E_{xv_1} \\ E_{v_1x} & E_{v_1v_1} \end{bmatrix} \in \mathbf{R}^{2(m+q_1)\times 2(m+q_1)}$$

has a triangular-time-Gaussian shape. As a result, matrices E_{xv_1} *and* $E_{v_1v_1}$ *are determined as blocks of* E_{uu}*. Then it follows that* $E_{v_1y} = E_{v_1x}A^T$ *and* $E_{yy} = AE_{xx}A^T + \sigma^2 I$*. Then* $G_0, H_0, \ldots, G_p, H_p$ *are determined by (3.6). In this case, condition (3.1) is true, i.e.* $\mathbf{v}_1$ *is well-defined.*

We wish to compare the accuracy associated with $\mathcal{T}_p(\mathbf{v}_0, \ldots, \mathbf{v}_p)$ *where* $G_0, H_0, \ldots, G_p, H_p$ *are determined by (3.6), for* $p = 0$ *and* $p = 1$*. We also wish to compare it with the accuracy associated with the method considered in [83] and called GBT2.*

In Figure 3.2, for $m = 10$ *and* $b_k = 0, 4$ *diagrams of the mean square error (MSE) associated with* $\mathcal{T}_0(\mathbf{v}_0)$*,* $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$ *and GBT2 are represented as follows. Figure 3.2 (a): the errors versus dimension* q_1 *of injection* $\mathbf{v}_1$*. Figure 3.2 (b): the errors versus parameter* σ *in noise* $\boldsymbol{\xi}$*. Figure 3.2 (c): the errors versus the reduction ratio.*

It follows from Figure 3.2 that the error associated with $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$ *is less than those of* $\mathcal{T}_0(\mathbf{v}_0)$ *and GBT2. In particular, in Figure 3.1 (a), the diagram of MSE associated with* $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$ *illustrates Theorems 4 and 6. First, the MSE decreases as* q *increases as Theorem 4 states. Second, for* $q \geq 30$*, condition (3.28) of Theorem 6 is true and the MSE associated with* $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$ *is less than those of GBT2. For* $q < 30$*, the MSE associated with* $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$ *is slightly equal than that of GBT2. This is because, for* $q < 30$*, condition (3.28) is not satisfied.*

3.5. Optimal model of the system with no reduction of input dimensionality

An important case of problem (2.6)-(2.8) is when the matrix product G_jH_j, for $j = 0, \ldots, p$, is replaced with a full rank matrix $P_j \in R^{m\times q_j}$. Then the input-output map $\mathcal{T}_h$ given by

$$\mathcal{T}_h(\mathbf{v}_0, \ldots \mathbf{v}_p) = \sum_{k=0}^{h} P_{kk} \tag{3.32}$$

is called *the full rank* system model of degree h. In Theorems 7 and 8 below, the optimal full rank model of degree h is determined and justified as a solution of problem (3.33). In problem (3.33) no constraint similar to that in (2.7) is imposed.

Theorem 7. *Let vectors $\mathbf{z}_0, \ldots, \mathbf{z}_p$ be pairwise uncorrelated. Then the minimal Frobenius norm solution to the problem*

$$\min_{P_0,\ldots,P_h} \|\mathcal{F}(\mathbf{y}) - \sum_{k=0}^{h} P_{kk}\|_\Omega^2, \tag{3.33}$$

is given, for $k = 0, \ldots, h$, by

$$P_k = E_{xz_k} E_{z_k z_k}^\dagger. \tag{3.34}$$

Proof. Similar to (3.8),

$$\|\mathcal{F}(\mathbf{y}) - \sum_{j=0}^{h} P_{jj}\|_\Omega^2 = \sum_{j=0}^{h} \|\mathbf{x} - P_j \mathbf{z_j}\|_\Omega^2 - tr\{h E_{xx}\}. \tag{3.35}$$

The minimal Frobenius norm solution to the problem

$$\min_{P_j} \|\mathbf{x} - P_j \mathbf{z_j}\|_\Omega^2$$

is given by (3.34). □

Theorem 8. *Let $A_j = E_{xz_j}(E_{z_j z_j}^{1/2})^\dagger$. The error associated with the minimal Frobenius norm solution to the problem in (3.33) is represented by*

$$\min_{P_0,\ldots,P_h} \|\mathcal{F}(\mathbf{y}) - \sum_{j=0}^{h} P_{jj}\|_\Omega^2 = \mathrm{tr}\{E_{xx}\} - \sum_{j=0}^{h} \|A_j\|^2, \tag{3.36}$$

where $\|\cdot\|$ is the Frobenius norm.

Proof. For P_j determined by (3.34),

$$\|\mathcal{F}(\mathbf{y}) - P_{jj}\|_\Omega^2 = tr\{E_{xx}\} - \|E_{xz_j}(E_{z_j z_j}^{1/2})^\dagger\|^2 + \|P_j E_{z_j z_j}^{1/2} - E_{xz_j}(E_{z_j z_j}^{1/2})^\dagger\|^2.$$

Here,

$$\begin{aligned} \|P_j E_{z_j z_j}^{1/2} - E_{xz_j}(E_{z_j z_j}^{1/2})^\dagger\|^2 &= \|E_{xz_j} E_{z_j z_j}^\dagger E_{z_j z_j}^{1/2} - E_{xz_j}(E_{z_j z_j}^{1/2})^\dagger\|^2 \\ &= \|E_{xz_j}(E_{z_j z_j}^{1/2})^\dagger - E_{xz_j}(E_{z_j z_j}^{1/2})^\dagger\|^2 \\ &= 0. \end{aligned}$$

That is,

$$\|\mathcal{F}(\mathbf{y}) - P_{jj}\|_\Omega^2 = tr\{E_{xx}\} - \|E_{xz_j}(E_{z_j z_j}^{1/2})^\dagger\|^2. \tag{3.38}$$

Then (3.36) follows from (3.8) and (3.38). □

Remark 9. *It follows from Theorem 8 that the error associated with the system model determined by (3.32) and Theorem 7 decreases if degree h increases.*

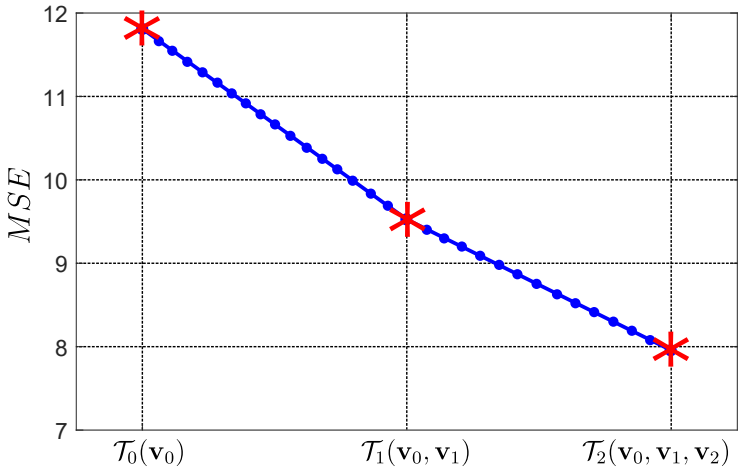


Figure 3.3: Example 4: Diagrams of the errors associated with $\mathcal{T}_0(\mathbf{v}_0)$, $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$ and $\mathcal{T}_2(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$.

Example 4. *Here, we wish to numerically illustrate Remark 9 and Theorem 8. To this end, we consider* $\mathbf{y} = \mathbf{x} + \boldsymbol{\xi}$ *where* $\mathbf{x} \in L^2(\Omega, R^m)$ *and* $\boldsymbol{\xi} \in L^2(\Omega, R^m)$ *are uniformly and normally distributed random vectors, respectively, with* $m = 50$. *Injections* $\mathbf{v}_1 \in L^2(\Omega, R^{q_1})$ *and* $\mathbf{v}_2 \in L^2(\Omega, R^{q_2})$ *are here chosen as uniformly distributed random vectors with* $q_1 = q_2 = m = 50$. *Noise* $\boldsymbol{\xi}$ *is uncorrelated with* $\mathbf{x}$ *and* $\mathbf{v_1}, \mathbf{v_2}$. *The mean square errors (MSE) associated with* $\mathcal{T}_0(\mathbf{v}_0)$, $\mathcal{T}_1(\mathbf{v}_0, \mathbf{v}_1)$ *and* $\mathcal{T}_2(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$ *are* $11.81, 9.52$ *and* 7.95, *respectively. They are diagrammatically shown in Figure 3.3, and illustrate Remark 9 and Theorem 8, i.e., the error associated with model* $\mathcal{T}_h$ *decreases if degree* h *increases.*

4. Conclusion and open problem

We have proposed and justified a new approach to the optimal modeling of non-linear system $\mathcal{F}$. It is assumed that its random input $\mathbf{y}$ and random output $\mathbf{x}$ are available and known in terms covariance matrices, and the input-output map $\mathcal{F}$ is unknown. The basic idea is to build the system model $\mathcal{I}_p$ as a sum of $p+1$ particular parts. The constructive determination of each part of the model is provided. The proposed model is optimal in the sense of minimization of the associated error. A significant ingredient of the proposed model is a set of so called injections $\mathbf{v}_1, \ldots, \mathbf{v}_p$ which allows us to further improve the accuracy of the system representation. The proposed model has several degrees of freedom to minimize the associated error.

An open associated problem is as follows: Given $\mathbf{x}, \mathbf{y}$ and $r_0, \ldots, r_p$, find matrices $G_0, H_0, \ldots, G_p, H_p$, injections $\mathbf{v}_1, \ldots, \mathbf{v}_p$ and transformations $Q_0, \ldots, Q_p$ that solve

$$\min_{\mathbf{v}_1,\ldots,\mathbf{v}_p} \min_{G_0,H_0,\ldots,G_p,H_p} \left\| \mathcal{F}(\mathbf{y}) - \sum_{j=0}^{p} G_j H_j \mathbf{z}_j \right\|_{\Omega}^2 \tag{4.1}$$

subject to

$$G_j \in R^{m\times r_j} \text{ and } H_j \in R^{r_j \times q_j}, \tag{4.2}$$

and

$$E_{z_i z_j} = \mathbf{O}, \quad \text{for } i \neq j, \tag{4.3}$$

where $i, j = 0, \ldots, p$.

An important difference from the problem in (2.6), (2.7) and (2.8) is that the problem in (4.1), (4.2) and (4.3) has p more unknowns to determine, they are $\mathbf{v}_1, \ldots, \mathbf{v}_p$. It makes this problem quite difficult. We are planning to tackle to it in the near future. A possible solution device is based on the special iterative procedure such that, for each iteration loop, the best approximation problem is solved providing searched optimal matrices $G_0, H_0, \ldots, G_p, H_p$ and optimal injections $\mathbf{v}_1, \ldots, \mathbf{v}_p$ with the smallest associated error.

References

[1] T. Adali and P. J. Schreier and L. L. Scharf, IEEE Transactions on Signal Processing, Complex-Valued Signal Processing: The Proper Way to Deal With Impropriety, Vol. 59 (11), pp. 5101-5125, 2011.

[2] 2009 Adamczak, Radoslaw and Litvak, Alexander E. and Pajor, Alain and Tomczak-Jaegermann, Nicole, Quantitative estimates of the convergence of the empirical covariance matrix in log-concave ensembles, Journal of the American Mathematical Society, Vol. 23, pp. 535-561, 2009.

[3] Orly Alter and Gene H. Golub, Singular value decomposition of genome-scale mRNA lengths distribution reveals asymmetry in RNA gel electrophoresis band broadening, Processing of the National Academy of Sciences of USA, Vol. 103 (32), pp. 11828-11833, 2006.

[4] Amar, A and Leshem, A and Gastpar, M., Recursive Implementation of the Distributed Karhunen-Loeve Transform, IEEE Transactions on Signal Processing, Vol 58 (10), pp. 5320-5330, 2010.

[5] D. Ambrosi and A. Quarteroni and G. Rozza (Eds)., Modelling of Physiological Flows, Vol. V, Springer-Verlag Italia, Milano, 2011.

[6] Zahra Aminzare and Eduardo D. Sontag, IEEE Tranactions on Network Science and Engineering, Synchronization of Diffusively-Connected Nonlinear Systems: Results Based on Contractions with Respect to General Norms, Vol.1 (2), pp. 91-106, 2014.

[7] Generalized Inverses Estimations by Means of Iterative Methods with Memory, Artidiello, S. and Cordero, A. and Torregrosa, J. R. and Vassileva, M., Mathematics, Multidisciplinary Digital Publishing Institute, Vol. 8 (1), pp. 2, 2020.

[8] Rasool Azimia and Mohadeseh Ghayekhlooa and Mahmoud Ghofranib and Hedieh Sajedic, Expert Systems with Applications, A novel clustering algorithm based on data transformation approaches, Vol. 76 (15), pp. 59-70, 2017.

[9] L. Bai and S. Dou and Z. Xiao and J. Choi, IET Signal Processing, Doubly iterative multiple-input-multiple-output-bit-interleaved coded modulation receiver with joint channel estimation and randomised sampling detection, Vol. 10 (4), pp. 335-341, 2016.

[10] A. Beck and L. Tetruashvili, SIAM J. Optim., On the Convergence of Block Coordinate Descent Type Methods, Vol. 23 (4), pp. 2037-2060, 2013.

[11] A. Ben-Israel and T. N. E. Greville, Generalized Inverses: Theory and Applications (2nd ed.), Springer, New York, 2003.

[12] Christos Boutsidis and David P. Woodruff, Optimal CUR Matrix Decompositions, SIAM Journal on Computing, Vol. 46 (2), pp. 543-589, 2017.

[13] S. Boyd and N. Parikh and E. Chu and B. Peleato and J. Eckstein, Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers, Foundations and Trends in Machine Learning, Vol. 3 (1), pp. 1-122, 2011.

[14] Fast low-rank modifications of the thin singular value decomposition, Brand, M., Linear algebra and its applications, Elsevier, Vol. 415 (1), pp. 20-30, 2006.

[15] Brillinger, David R., Time Series: Data Analysis and Theory, San Francisco, 2001.

[16] P. Bühlmann and S. Van de Geer, Statistics for High-Dimensional Data, Springer, New York, 2011.

[17] Digital Image Processing-An Algorithmic Introduction Using Java, Mark J. Burge and Wilhelm Burger, Springer-Verlag London, 2006.

[18] A family of higher-order convergent iterative methods for computing the Moore-Penrose inverse, Chen, H. and Wang, Y., Applied Mathematics and Computation, Elsevier, Vol. 218 (8), pp. 4012-4016, 2011.

[19] Tianping Chen and Hong Chen, IEEE Transactions on Neural Networks, Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems, Vol. 6 (4), pp. 911-917, 1995.

[20] S. Chenn and S. A. Billings, International Journal of Control, Representations of non-linear systems: the NARMAX model, 1989, Vol. 49 (3), pp. 1013-1032, 1989.

[21] S. Chenn and S. A. Billings, International Journal of Control, Neural networks for nonlinear dynamic system modelling and identification, Vol. 56 (2), pp. 319-346, 1992.

[22] S. Chenn and S. A. Billings and P. M. Grant, International Journal of Control, Non-linear system identification using neural networks, Vol. 51 (6), pp. 1191-1214, 1990.

[23] E. P. de Moura and F. de Abreu Melo Junior and F. Damasceno and L. Figueiredo and C. de Andrade and M. de Almeida, Renewable Energy, Classification of imbalance levels in a scaled wind turbine through detrended fluctuation analysis of vibration signals, Vol. 96, Part A, pp. 993-1002, 2016.

[24] Kamaludin Dingle and Chico Q. Camargo and Ard A. Louis, Nature Communications, Input-output maps are strongly biased towards simple outputs, Vol. 9 (761), 2018.

[25] Dragotti, P. L. and Gastpar, M., Distributed Source Coding: Theory, Algorithms and Applications, Academic Press, 2009.

[26] Maria Gabriela Eberle and Maria Cristina Maciel, Finding the closest Toeplitz matrix, Comput. Appl. Math., Vol. 22 (1), pp. 1-15, 2003.

[27] Fang, Jun and Li, Hongbin, IEEE Transactions on Signal Processing, Optimal/Near-Optimal Dimensionality Reduction for Distributed Estimation in Homogeneous and Certain Inhomogeneous Scenarios, Vol. 58 (8), pp. 4339-4353, 2010. doi 10.1109/TSP.2010.2048213.

[28] Vladimir N. Fomin and Michael V. Ruzhansky, SIAM Journal on Control and Optimization, Abstract Optimal Linear Filtering, Vol. 38 (5), pp. 1334-1352, 2000.

[29] L. Formaggia and A. Quarteroni and A. Veneziani (Eds.), Cardiovascular Mathematics-Modeling and simulation of the circulatory system, Vol. 1, Springer-Verlag Italia, Milano, 2009

[30] Friedland, S and Torokhti, A Generalized Rank-Constrained Matrix Approximations, SIAM Journal on Matrix Analysis and Applications, Vol. 29 (2), pp. 656-659, 2007

[31] Philip G. Gallman and K. S. Narendra, Automatica, Representations of nonlinear systems via the Stone-Weierstrass theorem, Vol. 12 (6), pp. 619-622, 1976.

[32] Gianfelici, Francesco and Turchetti, Claudio and Crippa, Paolo, A non-probabilistic recognizer of stochastic signals based on KLT, Signal Processing, Vol. 89 (4), pp. 422-437, 2009.

[33] F. Gini and F. Lombardini, IEEE Aerospace and Electronic Systems Magazine, Multibaseline cross-track SAR interferometry: a signal processing perspective, Vol. 20 (8), pp. 71-93, 2005.

[34] P. G. Howlett and A. P. Torokhti and Pearce, C. E. M., A Philosophy for the Modelling of Realistic Non-linear Systems, Processing of Amer. Math. Soc., Vol. 131 (2), pp. 353-363, 2003.

[35] Hua, Y. and Wanquan Liu, IEEE Signal Processing Letters, Generalized Karhunen-Loeve transform, Vol. 5 (6), pp. 141-142, 1998.

[36] Hua, Y. and Nikpour, M. and Stoica, Petre, Optimal reduced-rank estimation and filtering, IEEE Transactions on Signal Processing, Vol. 49 (3), pp. 457-469, 2001.

[37] Ian T. Jolliffe and Jorge Cadima, Principal component analysis: a review and recent developments, Phil. Trans. R. Soc. A., Vol. 374, pp. 1-16, 2016.

[38] I. T. Jolliffe, Principal Component Analysis, Second Edition, Springer, New York, 2002.

[39] Joong-Ho Won, Johan Lim, Seung-Jean Kim and Bala Rajaratnam, Condition-number-regularized covariance estimation, Journal of the Royal Statistical Society: Series B), Vol. 75 (3), pp. 427-450, 2013.

[40] T. Kailath and A. Saed and B. Hassibi, Linear estimation, Prentice-Hall, Upper Saddle River, 2000.

[41] Kamrani, Hossein and Asli, Alireza Zolghadr and Markopoulos, Panos P. and Langberg, Michael and Pados, Dimitris A. and Karystinos, George N., Reduced-Rank L1-Norm Principal-Component Analysis With Performance Guarantees, IEEE Transactions on Signal Processing, Vol. 69, pp. 240-255, 2021. doi 10.1109/TSP.2020.3039599.

[42] Vassilis Kekatos and Georgios B. Giannakis, IEEE Transactions on Signal Processing, Sparse Volterra and Polynomial Regression Models: Recoverability and Estimation, Vol. 59 (12), pp. 5907-5920, 2011.

[43] Kitagawa, G. and Gersch, W., Estimation of Time Varying Variance, Smoothness Priors Analysis of Time Series, Springer New York, pp. 137-145, 1996.

[44] Roumen Kmuntchav and Roumiena Kountcheva, Hierarchical Adaptive KL-Based Transform: Algorithms and Applications, Computer Vision in Control Systems-1, Favorskaya, Margarita N and Jain, Lakhmi C. (Eds.), Vol. 73, pp. 91-136, Springer, 2015.

[45] Lara, M. and Mulgrew, B., Proceedings of the 20th European Signal Processing Conference (EUSIPCO), Performance of the distributed KLT and its approximate implementation, pp. 724-728, 2012.

[46] Ledoit, Olivier and Wolf, Michael, A well-conditioned estimator for large-dimensional covariance matrices, Journal of Multivariate Analysis, Vol. 88 (2), pp. 365-411, 2004.

[47] Ledoit, Olivier and Wolf, Michael, Nonlinear shrinkage estimation of large-dimensional covariance matrices, The Annals of Statistics, The Institute of Mathematical Statistics, Vol. 40 (2), pp. 1024-1060, 2012.

[48] Junlin Li and Ghassan AlRegib, Distributed Estimation in Energy-Constrained Wireless Sensor Networks, IEEE Transactions on Signal Processing, Vol. 57 (10), pp. 3746-3758, 2009.

[49] Y. Liao and K. Chen, IEEE Transactions on Wireless Comm., Estimation of Stationary Phase Noise by the Autocorrelation of the ICI Weighting Function in OFDM Systems, Vol. 5 (12), pp. 3370-3374, 2006.

[50] Krylov Subspace Methods: Principles and Analysis, Liesen, J. and Strakos, Z., OUP Oxford, 2013.

[51] Wei Liu and Hongwei Zhang and Kaijiang Yu and Xingguo Tan, Optimal linear filtering for networked systems with communication constraints, fading measurements, and multiplicative noises, International Journal of Adaptive Control and Signal Processing, 2016. https://onlinelibrary.wiley.com/doi/abs/10.1002/acs.2745.

[52] Xifu Liu and Wen Li and Hongxing Wang, Rank constrained matrix best approximation problem with respect to (skew) Hermitian matrices, Journal of Computational and Applied Mathematics, Vol. 319, pp. 77-86, 2017.

[53] Lombardini, Fabrizio and Gini, Fulvio, Model Order Selection in Multi-baseline Interferometric Radar Systems, EURASIP Journal on Advances in Signal Processing, Vol. 20, pp. 3206-3219, 2005.

[54] Lu, Canyi and Feng, Jiashi and Lin, Zhouchen and Mei, Tao and Yan, Shuicheng, IEEE Transactions on Pattern Analysis and Machine Intelligence, Subspace Clustering by Block Diagonal Representation, Vol. 41 (2), pp. 487-501, 2019, doi 10.1109/TPAMI.2018.2794348.

[55] Effective algorithms of the Moore-Penrose inverse matrices for extreme learning machine, Lu, S. and Wang, X. and Zhang, G. and Zhou, X., Intelligent Data Analysis, IOS Press, Vol. 19 (4), pp. 743-760, 2015

[56] Hang Ma and Yu-Han Yang and Yan Chen and K. J. Ray Liu and Qi Wang, Distributed State Estimation With Dimension Reduction PreProcessingIEEE Transactions on Signal Processing, Vol. 62 (12), pp. 3098-3110, 2014.

[57] Damian Edgardo Marelli and Minyue Fu, Distributed weighted least-squares estimation with fast convergence for large-scale systems, Automatica, Vol. 51, pp. 27-39, 2015.

[58] J. Navarro-Moreno and M. D. Estudillo-Martinez and R. M. Fernandez-Alcala and J. C. Ruiz-Molina, IEEE Transactions on Information Theory, Estimation of Improper Complex-Valued Random Signals in Colored Noise by Using the Hilbert Space Theory, Vol. 55, pp. 2859-2867, 2009.

[59] Papadatos, N. and Xifara, T., A simple method for obtaining the maximal correlation coefficient and related characterizations, Journal of Multivariate Analysis, Vol. 118, pp. 102-114, 2013.

[60] Phophalia, Ashish and Rajwade, Ajit and K. Mitra, Suman, Rough set based image denoising for brain MR images, Signal Processing, Vol. 103, pp. 24-35, 2014.

[61] Stochastic MV-PURE T. Piotrowski and R. Cavalcante and I. Yamada, Estimator-Robust Reduced-Rank Estimator for Stochastic Linear Model, IEEE Transactions on Signal Processing, Vol. 57 (4), pp. 1293-1303, 2009.

[62] Tomasz Piotrowski and Isao Yamada, Performance of the stochastic MV-PURE estimator in highly noisy settings, Journal of The Franklin Insitute, Vol. 351, pp. 3339-3350, 2014.

[63] L. Piroddi and W. Spinelli, International Journal of Control, An identification algorithm for polynomial NARX models based on simulation error minimization, Vol. 76 (17), pp. 1767-1781, 2003.

[64] Foundations of Mathematical Analysis, Ponnusamy, S., Birkhäuser Boston, 2011.

[65] H. V. Poor, An Introduction to Signal Processing and Estimation, Springer-Verlag, 2nd ed. New York., 2001.

[66] Saghri, J. A. and Schroeder, S. and Tescher, A. G., Adaptive two-stage Karhunen-Loeve-transform scheme for spectral decorrelation in hyperspectral bandwidth compression, Optical Engineering, Vol. 49 (5), pp. 1-7, 2010.

[67] I. W. Sandberg, IEEE Trans. on Circuits and Syst. Part 1. Fund. Theory and Applic., Notes on uniform approximation of time-varying systems on finite time intervals, Vol. 45, pp. 863-864, 1998.

[68] I. W. Sandberg, IEEE Trans. on Circuits and Syst. Part 1. Fund. Theory and Applic., R+ Fading Memory and Extensions of Input-Output Maps, Vol. 49 (11), pp. 1586-1591, 2002.

[69] I. W. Sandberg, Time-Delay Polynomial Networks and Quality of Approximation, IEEE Trans. on Circuits and Systems. Part 1, Fundamental theory and applications, Vol. 47, pp. 40-49, 2000.

[70] Scharf, L., The SVD and reduced rank signal processing, Signal Processing, Vol. 25 (2), pp. 113-133, 1991.

[71] Schizas, I. D. and Giannakis, G. B. and Zhi-Quan Luo, IEEE Transactions on Signal Processing, Distributed Estimation Using Reduced-Dimensionality Sensor Observations, Vol. 55 (8), pp. 4284-4299, 2007.

[72] Michael K. Schneider and Alan S. Willsky, A Krylov Subspace Method for Covariance Approximation and Simulation of Random Processes and Fields, Multidimensional Systems and Signal Processing, Vol. 14 (4), pp. 295-318, 2003.

[73] P. J. Schreier and L. L. Scharf and C. T. Mullis, Detection and estimation of improper complex random signals, IEEE Transactions on Information Theory, Vol. 51 (1), pp. 306-312, 2005.

[74] Yiyuan She and Shijie Li and Dapeng Wu, Robust Orthogonal Complement Principal Component Analysis, Journal of the American Statistical Association, Vol. 111 (514), pp. 763-771, 2016.

[75] Song, Enbin and Zhu, Yunmin and Zhou, Jie, Sensors' optimal dimensionality compression matrix in estimation fusion, Automatica, Vol. 41 (12), pp. 2131-2139, 2005.

[76] Stein, C., Processing Prague Symp. Asymptotic Statist., Estimation of the mean of a multivariate normal distribution, pp. 345-381, 1973.

[77] Michael Stewart, A Superfast Toeplitz Solver with Improved Numerical Stability, SIAM Journal on Matrix Analysis and Applications, Vol. 25 (3), pp. 669-693, 2003.

[78] Strobach, P., New forms of Levinson and Schur algorithms, IEEE Signal Processing Magazine, Vol. 8 (1), pp. 12-36, 1991. doi 10.1109/79.80993.

[79] V. N. Temlyakov, Nonlinear Methods of Approximation, Foundations of Computational Mathematics, Vol. 3 (1), pp. 33-107, 2003.

[80] A. Torokhti and P. Howlett, Computational Methods for Modelling of Nonlinear Systems, Elsevier, Amsterdam, 2007.

[81] Torokhti, A. and Friedland, S., Towards theory of generic Principal Component Analysis, Journal of Multivariate Analysis, Vol. 100 (4), pp. 661-669, 2009.

[82] Anatoli Torokhti and Phil Howlett and Hamid Laga, Estimation of stochastic signals under partially missing information, Signall Processing, Vol. 11, pp. 199-209, 2015.

[83] Anatoli Torokhti and Pablo Soto-Quiros, Generalized Brillinger-Like Transforms, IEEE Signal Processing Letters, pp. 843-847, Vol. 23 (6), 2016.

[84] Tseng, P., Convergence of a Block Coordinate Descent Method for Nondifferentiable Minimization, Journal of Optimization Theory and Applications, Kluwer Academic Publishers-Plenum Publishers, Vol. 109 (3), pp. 475-494, 2001.

[85] Vershynin, Roman, J. Th. Prob., How Close is the Sample Covariance Matrix to the Actual Covariance Matrix?, Springer US, Sample covariance matrices; Estimation of covariance matrices; Random matrices with independent columns; 60H12; 60B20; 46B09, Vol. 25 (3), pp. 655-686, 2012.

[86] Rene Vidal and Yi Ma and Shankar Sastry, Generalized Principal Component Analysis, Springer-Verlag New York, 2016.

[87] H. Wang, Rank constrained matrix best approximation problem, Applied Mathematics Letters, Vol. 50, pp. 98-104, 2015. issn 0893-9659.

[88] Wang, Qian and Jing, Yindi, New rank detection methods for reduced-rank MIMO systems, EURASIP Journal on Wireless Communications and Networking, Vol. 2015 (1), pp. 1-16, SpringerOpen, 2015.

[89] Songcen Xua and Rodrigo C. de Lamare and H. Vincent Poor, Signal Processing, Distributed low-rank adaptive estimation algorithms based on alternating optimization, Vol. 144, pp. 41-51, 2018.

[90] R. Yang and J. O. Berger, The Anna1s of Statistics, Estimation of a covariance matrix using the reference prior, Springer US, Sample covariance matrices; Estimation of covariance matrices; Random matrices with independent columns; 60H12; 60B20; 46B09, Vol. 22 (3), pp. 1195-1211, 1994.

[91] Linxia Zhang and Dunbiao Niu and Enbin Song and Jie Zhou and Qingjiang Shi and Yunmin Zhu, IEEE Transactions on Signal Processing, Joint Optimization of Dimension Assignment and Compression in Distributed Estimation Fusion, Vol. 67 (9), pp. 2453-2468, 2019.

[92] Xiao-Long Zhu and Xian-Da Zhang and Zi-Zhe Ding and Ying Jia, PCA algorithms for blind source separation without prewhitening, IEEE Transacvions on Circuits ind Systems-I, Vol. 53 (3), pp. 745-753, 2006.

[93] Yunmin Zhu and Enbin Song and Jie Zhou and Zhisheng You, IEEE Transactions on Signal Processing, Optimal dimensionality reduction of sensor data in multisensor estimation fusion, Vol. 53 (5), pp. 1631-1639, 2005.

Anatoli Torokhti

STEM discipline,

University of South Australia,

Mawson Lakes, SA 5076,

Australia

e-mail: anatoli.torokhti@unisa.edu.au

Corresponding author

and

Pablo Soto-Quiros

Escuela de Matemática,

Instituto Tecnológico de Costa Rica,

Cartago,

Costa Rica

e-mail: jusoto@tec.ac.cr